Methylation in host and viral genes as marker of aggressiveness in cervical lesions: Analysis in 543 unscreened women

Methylation in host and viral genes as marker of aggressiveness in cervical lesions: Analysis in 543 unscreened women

Valentina Fiano [a,*], Morena Trevisan [a], Francesca Fasanelli [a], Chiara Grasso [a], Federica Marabese [a], Maria da Graça Bicalho [b], Newton S. de Carvalho [c], Carlos A. Maestri [d], Franco Merletti [a,e], Carlotta Sacerdote [e], Laura De Marco [a,e,1], Anna Gillio-Tos [a,1]

[a] Unit of Cancer Epidemiology–CeRMS, Department of Medical Sciences, University of Turin, Via Santena 7, 10126 Turin, Italy
[b] Laboratory of Immunogenetics and Hystocompatibility (LIGH), Federal University of Paranà, Rua XV de Novembro, 1299, Curitiba, PR 80060-000, Brazil
[c] Department of Gynecology and Obstetrics, Federal University of Paraná, Infectious Diseases in Gynecology and Obstetrics Sector, Hospital de Clínicas, Rua XV de Novembro, 1299, Curitiba, PR 80060-000, Brazil
[d] Department of Cervical Pathology, Hospital Erasto Gaertner, Curitiba, R. Dr. Ovande do Amaral, 201 - Jardim das Americas, Curitiba, PR 81520-060, Brazil
[e] Unit of Cancer Epidemiology, Città della Salute e della Scienza University-Hospital and Center for Cancer Prevention (CPO), Via Santena 7, 10126 Turin, Italy

HIGHLIGHTS

- Methylation in CADM1 and MAL is associated with high grade cervical cancer lesions.
- Methylation in HPV regions is associated with high grade cervical cancer lesions.
- Methylation in HPV regions is associated with age.
- Increasing number of methylated genes predicts high grade cervical lesions.

abstract

Objective. The present study aimed to evaluate the association between altered methylation and histologically confirmed high grade cervical intraepithelial neoplasia (hgCIN).

Methods. Methylation levels in selected host (CADM1, MAL, DAPK1) and HPV (L1_I, L1_II, L2) genes were measured by pyrosequencing in DNA samples obtained from 543 women recruited in Curitiba (Brazil), 249 with hgCIN and 294 without cervical lesions. Association of methylation status with hgCIN was estimated by Odds Ratio (OR) with 95% confidence interval (CI).

Results. The mean methylation level increased with severity of the lesion in the host and viral genes (p-trend b 0.05), with the exception of L1_II region (p-trend = 0.075). Positive association was found between methylation levels for host genes and CIN2 and CIN3 lesions respectively [CADM1: OR 4.17 (95% CI 2.03-8.56) and OR 9.54 (95% CI 4.80-18.97); MAL: OR 5.98 (95% CI 2.26-15.78) and OR 22.66 (95% CI 9.21-55.76); DAPK1: OR 3.37 (95% CI 0.93-12.13) and OR 6.74 (95% CI 1.92-23.64)]. Stronger risk estimates were found for viral genes [L1_I: OR 10.74 (95% CI 2.66-43.31) and OR 15.00 (95% CI 3.00-74.98); L1_II: OR 73.18 (95% CI 4.07-1315.94) and OR 32.50 (95% CI 3.86-273.65); L2: OR 4.73 (95% CI 1.55-14.44) and OR 10.62 (95% CI 2.60-43.39)].
The cumulative effect of the increasing number of host and viral methylated genes was associated with the risk of CIN2 and CIN3 lesions (p-trend b 0.001).

Conclusions. Our results, empowered by a wide cervical sample series with a large number of hgCIN, sup- ported the role of methylation as marker of aggressiveness.

* Corresponding author at: Cancer Epidemiology – CeRMS, University of Turin, via Santena 7, 10126 Turin, Italy.
E-mail addresses: valentina.fiano@unito.it (V. Fiano), morena.trevisan@unito.it (M. Trevisan), francesca.fasanelli@unito.it (F. Fasanelli), chgrasso@unito.it (C. Grasso), federica.marabese@edu.unito.it (F. Marabese), mgbicalho@gmail.com (M. da Graça Bicalho), newtonsdc@gmail.com (N.S. de Carvalho), camcmi@uol.com.br (C.A. Maestri), franco.merletti@unito.it (F. Merletti), carlotta.sacerdote@cpo.it (C. Sacerdote), laura_demarco@virgilio.it (L. De Marco), gilliotos.demarco@cpo.it (A. Gillio-Tos).
[1] These authors contributed equally to this work.

1. Introduction

Incidence and mortality of cervical cancer have been reduced in the last decades in industrialized countries by organized screening pro- grams. Nevertheless, cervical cancer still remains a public health issue affected by several unresolved aspects of the natural history of the dis- ease. Despite the identification of the etiological role of the human pap- illomavirus (HPV), the events that promote or prevent viral persistence in cervical cells, as well as the determinants of cervical cancer progres- sion, have not been yet identified.
DNA methylation at CpG sites, mostly affecting gene expression, has been described as associated with carcinogenesis in many cancer sites [1,2]. Altered methylation in host genes and in HPV genome has been recently reported in cervical cancer [3–10] and suggested as indicator of transforming HPV infection and potential biomarker of aggressive- ness. However studies on the association between the level of DNA methylation and the severity of the cervical lesions still led to inconclu- sive results [7,8], specifically considering that the strength of association may depend on selection of the investigated CpG sites and gene regions. Among the cell genes investigated for methylation status in their promoter, CADM1 (Cell Adhesion Molecule 1), MAL (T-Lymphocyte Maturation protein) and DAPK1 (Death Associated protein kinase 1) gave more consistent results across the studies [6]. These genes are directly involved in crucial cell pathways and their silencing could have biologically plausible effects on cell cycle de-regulation. The HPV DNA includes a variable number of CpG sites, depending on different ge- notypes, spread out in the viral genome and poorly methylated [11]. They may thereafter undergo methylation by the host cell methylation machinery. Measuring DNA methylation in HPV genome has shown promise for accurate detection of high grade cervical intraepithelial le- sions (hgCIN) [7,9,10,12–17]. Several studies investigated the methyla- tion of specific viral genes as well as the overall proportion of CpGs methylated in the whole viral genome. However they were mostly fo- cused on HPV16, encountering limited sample size, showing different frequencies of methylation at specific sites associated to high grade le- sions and leading to results somehow not conclusive. A study nested in the Guanacaste cohort highlighted a cross-sectional and longitudinal association of cervical pre-cancer lesion with methylation in HPV16 L1, L2 and E2 genes. Stronger effects were found when the analyses were restricted to specific CpGs: L1 nucleotide positions 5601-5616 and 6457; L2 nucleotide position 4261 [14,15]. Viral types different from HPV16 were investigated at a lesser extent [12,13,15,18–21].

The suggested role of host and viral methylation in the natural his- tory of progression of HPV infected cells to hgCIN and cancer still needs confirmation by studies with wide sample size and enlarged to genotypes different from HPV16.

In the current study we evaluated the association between altered methylation and histologically confirmed hgCIN in a wide cervical sam- ple series

obtained from Brazilian women originally recruited in the frame of awareness campaigns for adhesion to cervical screening, as in Brazil organized screening programs are not present.

We investigated the association of the methylation levels in selected host (CADM1, MAL and DAPK1) and viral (L1 and L2 of HPV 16, 18, 31, 45) genes with the severity of the cervical lesions.

2. Methods

2.1. Study design and participants

A cross-sectional study was originally set up including samples from subjects recruited as previously described [22] within a collaboration among the Unit of Cancer Epidemiology (Turin, Italy), the Laboratory of Immunogenetics and Histocompatibility (LIGH), the Department of Gynecology and Obstetrics, at the Federal University of Paraná, Infec- tious Diseases in Gynecology and Obstetrics Sector and the Department of Cervical Pathology, Hospital Erasto Gaertner (Curitiba, Brazil). Briefly, under supervision of LIGH, women aged 15 to 47 years were recruited in Curitiba, between February and June 2010 by local gynecologists through awareness campaigns for adhesion to cervical screening, as in Brazil or- ganized screening programs are not present. Pregnant women were ex- cluded. Women were offered cervical cytology with Papanicolaou staining, further colposcopy if needed with diagnostic criteria (Bethesda 2001 system) previously described [22]. They were also asked to donate a cervical sample collected in sample transport medium (STM, Qiagen, Hilden, Germany) for HPV testing and further molecular investigations. Approval was released by the Ethical Committee for Clinical Research of the Hospital Erasto Gaertner (protocol CEP: 81520-060, P.P No 1943). At enrolment all participating women were administered a ques- tionnaire by interview, which collected demographic, sexual and lifestyle information including age, ethnic group, lifetime number of sexual part- ners and smoking status.

Within this frame, we selected overall 543 women, 249 with histo- logically confirmed hgCIN (137 CIN2 and 112 CIN3) defined as cases and 294 women without lesions defined as controls. In the aim to eval- uate if altered methylation can be associated to hgCIN, women with CIN1 lesions were excluded, as CIN1 is widely proven to have a low rate of neoplastic progression to high grade. Furthermore we excluded from the study controls originally included in the population if DNA was not available or not sufficient for methylation analysis.

The DNA had been extracted from cervical cells collected in STM at the LIGH and stored at −80 °C at Unit of Cancer Epidemiology (Turin, Italy). DNA had been screened for HPV infection and genotyped if posi- tive. Details on genomic DNA extraction, HPV detection (GP5+/6+ PCR) and HPV genotyping by multiplex PCR were previously described [22]. Methylation analyses were performed for host genes on all the study samples, while for viral genes they were restricted to samples (N = 160) with single type infection by HPV16 (N = 122), 18 (N = 12), 31 (N = 21) and 45 (N = 5).

2.2. Genes and CpG selection

Genes were selected relying on consistency across the recent litera- ture of results which provided association between altered methylation and hgCIN occurrence: CADM1, MAL and DAPK1 among host genes, L1 and L2 among HPV genes were chosen. Target CpGs for host genes (Supplementary Table S1) were identified in the promoter according to pub- lished reports [12,16].

For viral genes the HPV16 genome sequence was used as reference. L1 and L2 target CpG sites (Supplementary Table S1) were identified according to published reports [12,14,15,19,20] and selected as reported significantly associated with hgCIN and/or with OR ≥ 2 in at least two studies. We targeted two regions in the L1 viral gene hereafter named as L1 I and L1 II. For HPV18, 31, 45, the target CpGs were identified after alignment with HPV16 sequence through the multiple sequence alignment tool Clustal Omega (www.ebi.ac.uk/Tools/msa/clustalo/), which allowed to visualize the sequence homology and identify for these types the most nearby CpG to each selected HPV16 CpG. Sequence alignment was performed following conversion to bisulfite modified se- quences by PyroMark Assay Design 2.0 software (Qiagen). The correct format to fit Clustal Omega tool was obtained by the Emboss Seqret soft- ware (www.ebi.ac.uk/Tools/sfc/emboss_seqret). Reference sequences and CpG positions in each target region for each HPV type were found at www.ncbi.nlm.nih.gov web site and are listed in the Supplementary Table S1.

2.3. Methylation assays

Primers for the selected target regions were designed by using PyroMark Assay Design 2.0 software (Qiagen) according to the software stringent criteria. The sequences of the optimized primers, along with their annealing profile, are listed in the Supplementary Table S2.

DNA samples (≤1 μg), as well as methylation controls (i.e. synthetic methylated and unmethylated DNA (Qiagen)), SiHa and HeLa cell line DNA, synthetic HPV16 and HPV18 plasmidic DNA containing the com- plete genomes (Medical System, Genoa, Italy) underwent bisulfite modification by the Epitect Bisulfite Kit (Qiagen) according to the man- ufacturer's instructions, except for the optimized incubation time ex- tended to 16 h [23]. Plasmidic DNA underwent two cycles of bisulfite modification to achieve a complete conversion.

Methylation assays were performed by pyrosequencing onto a PyroMark Q24 MDx system (Qiagen). Pyrosequencing assays (CpG mode, software Q24 version 2.0) were performed as previously described [24]. For the host genes preliminary PCR reactions were performed using PyroMark PCR kit (Qiagen) according to the manufacturer's instruction except for annealing temperature specific for each gene (Supplementary Table S2). For the viral regions L1 I and L1 II preliminary PCR reactions were performed in a total volume of 35 μl containing 1× PCR buffer, 2 mM $MgCl_2$, 0.8 mM dNTPs, 0.5 μM of each primer, 1.75 U Taq polymer- ase and 2 μl of bisulfte modified DNA. The cycling profile was as follows: 95 °C for 10 min followed by 45 cycles of denaturation at 95 °C for 30 s, annealing for 1 min at the specific temperature set for each HPV type as listed in Supplementary Table S2, extension at 72 °C for 1 min, final ex- tension at 72 °C for 10 min. For the L2 region PCRs were performed using the PyroMark PCR Kit (Qiagen) following manufacturer instruc- tions, except for the annealing temperature set for each hrHPV type as listed in Supplementary Table S2.

The methylation cut off were determined by considering the limit of the quantification (LoQ) of the pyrosequencing which can be calculated by the average methylation of multiple repeats of negative samples and

its standard deviation (SD) [25-27]. For a conservative approach we considered the addition of 10 × SD to the average value of unmethylated samples as suggested by Lehmann [27]. The methylation cut off for each selected host or viral gene was achieved through the evaluation of multiple repeats (N = 23) of samples (4 for each studied gene) with normal cytology and unmethylated in all the studied CpG sites.

2.4. Statistical methods

Descriptive statistics of participants were reported using means, standard deviations (SD), counts and percentages according to the cyto/histological category (HPV negative controls, HPV positive con- trols, CIN2/3, CIN2 and CIN3 cases). Differences between controls and CIN2/3 cases were tested using Mann Whitney test for continuous data and Chi Square or Fisher test for categorical data. Test for trend for all methylation variables was assessed using Cuzick nonparametric test (*nptrend* command, ordered groups: all normal/CIN2/CIN3).

Box plots of the mean methylation were produced for the studied categories, Spearman's rank correlation coefficients were used to quan- tify the correlation of the mean methylation values between host and viral genes and Scatter plots were drafted to graphically visualize the association.

Logistic regression models were used to estimate odds ratios (ORs), and corresponding confidence intervals (CIs) and p-values of CIN2/3, CIN2 and CIN3 for methylation values equal or above the determined cut off. A crude and an adjusted model for age (continue variable), eth- nic group (white versus other), smoking status (ever versus never) and number of sexual partners (N3 versus ≤3) were fitted, to address

Table 1
Characteristics of the study subjects by cyto/histological status.

	Controls		Cases			All	p-Value[c]
Characteristcs	Normal cytology		CIN2 + CIN3	CIN2	CIN3		
Total subjects	HPV− N = 200	HPV+ N = 94	N = 249	N = 137	N = 112	N = 543	
Mean age (SD)	32.61 (8.08)	28.71 (7.22)	31.64 (6.56)	31.07 (6.85)	32.35 (6.15)	31.49 (7.38)	p = 0.60
HPV status							
HPV+		94	249	137	112	343	
single infection		60	163	89	34	223	
multiple infection		28	83	46	37	111	
N.D.[a]		6	3	2	1	9	
HPV16		22	100	48	52	122	
HPV31 or 18 or 45		15	23	16	7	38	
Smoking status							
Never	111 (56%)	44 (47%)	117 (47%)	65 (48%)	52 (46%)	272 (50%)	
Ever	87 (44%)	50 (53%)	131 (53%)	71 (52%)	60 (54%)	268 (50%)	p = 0.171
Missing	2	0	1	1	0	3	
Sexual partners							
≤3	125 (62%)	42 (45%)	109 (44%)	59 (43%)	50 (45%)	276 (51%)	
N3	75 (38%)	52 (55%)	139 (56%)	77 (57%)	62 (55%)	266 (49%)	p = 0.03
Missing	0	0	1	1	0	1	
Ethnic group							
White	95 (47%)	36 (39%)	82 (33%)	39 (28%)	43 (38%)	213 (39%)	
Black	12 (6%)	3 (4%)	6 (2%)	5 (4%)	1 (1%)	21 (4%)	
Brasilian mixed[b]	64 (32%)	45 (48%)	143 (57%)	83 (61%)	60 (54%)	252 (47%)	
Asiatic	1 (1%)	0 (0%)	1 (1%)	0 (0%)	1 (1%)	2 (0%)	p b 0.001
Not available	28 (14%)	10 (11%)	17 (7%)	10 (7%)	7 (6%)	55 (10%)	
Mean percentages of methylation (SD)							p-Trend[d]
All subjects							
CADM1	2.04% (1.16)	2.49% (1.76)	4.15% (5.10)	3.50% (4.59)	4.94% (5.60)	3.08% (3.73)	p b 0.001
MAL	2.24% (0.75)	2.68% (1.46)	3.63% (2.96)	2.97% (2.11)	4.44% (3.58)	2.95% (2.24)	p b 0.001
DAPK1	1.69% (1.36)	2.22% (1.96)	2.97% (4.89)	2.40% (3.29)	3.91% (6.67)	2.33% (3.45)	p = 0.038
HPV16 or 18 or 31 or 45		N = 37	N = 123	N = 64	N = 59	N = 160	
L1_I		19.68% (25.10)	20.92% (21.15)	18.04% (18.89)	24.05% (23.13)	20.68% (21.88)	p = 0.010
L1_II		17.80% (23.92)	22.30% (21.77)	21.65% (18.12)	23.00% (25.28)	21.44% (22.18)	p = 0.075
L2		16.80% (24.23)	19.10% (16.70)	17.21% (14.59)	21.20% (18.68)	18.68% (18.25)	p = 0.016

[a] N.D. HPV type not identified by the genotyping system (Digene HPV genotyping RH, Qiagen).
[b] Brazilian Mixed: a miscegenation of Euro-Descendent, Afro-Descendent, Amerindian and East Asian [34].
[c] p-Value assessed comparing controls versus cases CIN2/3.
[d] p-Trend = p-value for trend assessed using Cuzick nonparametric test (*nptrend* command, ordered groups: controls/CIN2/CIN3).

potential source of bias. In particular for smoking status subjects were categorized as “never” if never smokers and as “ever” if ex or current smokers. Women with missing data (Supplementary Table S3) on age, smoking status and number of sexual partners were excluded in the ad- justed analysis. Women of unknown ethnicity were included in the “other” category of the variable “ethnic group”.

The analyses with CIN2/3 as outcome were further stratified by age considering the categories <35 and ≥35 (previously the interaction be- tween age and methylation was tested by including the cross-product interaction term in the model and then using the Wald test to assess the statistical significance). In the case of empty cells the logistic regres- sion model was fitted by penalized maximum likelihood regression.

The combined effect of methylation of host and viral genes was an- alyzed by creating composite variables (values 0 to 3) by number of genes found above the threshold, then used as exposure variables in a logistic model considering the outcome CIN2/3. Linearity of trends across the three created categories was tested by considering the cate- gorical variable as a continuous variable in the logistic model.

All statistical analyses were performed using STATA version 13 (STATA Corp., Texas, USA).

3. Results

3.1. Study population overview

The characteristics of cases (N = 249) and controls (N = 294) are reported in Table 1 and the missing data in Supplementary Table S3. Out of the 249 cases, 163 had infection by single HPV type, 83 by multi- ple HPV types, 3 by HPV types not detectable by the used genotyping kit. Out of 294 controls, 200 were negative and 94 positive to HPV infection. Among HPV positive controls 60 had infection by single HPV type, 28 by multiple HPV types, 6 by HPV types not detectable by the used genotyp- ing kit. A statistically significant difference between cases and controls in the ever smokers category was not found (p-value = 0.171, Table 1). The proportion of women with more than three sexual part- ners was 38% in the controls without HPV infection, 55% in the controls with HPV infection and 56% in the cases (p-value of the comparison con- trols versus cases = 0.030). Controls encountered a higher frequency of white ethnic group than cases (47% for HPV negative and 39% for HPV positive versus 33%) whereas lower frequency of Brazilian mixed (32% for HPV negative and 48% for HPV positive versus 57%) is present (p- value < 0.001).

The mean methylation level increased with severity of disease (Fig. 1). An increasing trend was seen for all the study genes and target regions with statistical significance except for L1 II as reported in Table 1.

Host and viral genes methylation and association with high grade cervical lesions.

The methylation cut off achieved through the evaluation of multiple repeats of controls samples with normal cytology and unmethylated in all the studied CpG sites was 9% for DAPK1 and 5% for all the other host genes and viral regions.

Methylation levels showed a low correlation in the host genes CADM1 and MAL (rho = 0.43, Bonferroni-corrected p-value = 0.0001) (Supplementary Fig. S1) and a moderate/high correlation in viral genes (rho = 0.64 between L1 I and L1 II, rho = 0.66 between L1 II and L2 and rho = 0.79 between L1 I and L2; Bonferroni-corrected p- values < 0.0001). There was no correlation between host and viral genes methylation levels.

Methylation in CADM1, MAL or DAPK1 genes was associated with an increased risk of hgCIN (Table 2). The association was statistically signif- icant for CADM1, MAL and DAPK1 overall and for CIN3. For CIN2, the as- sociation is significant for CADM1 and MAL but not for DAPK1. In the HPV positive restricted analyses, the association remained strong and statistically significant overall and for CIN3 for CADM1 and MAL (Table 2). The risk of hgCIN remained after adjustment for age, ethnic group, smoking status and number of sexual partners.

When considering viral genes, the methylation levels of all viral re- gions showed a strong and significant (p-values < 0.05) association with hgCIN, both in HPV positive women and in the HPV16 subgroup (Table 2), which remained after stratification for CIN2 and CIN3 lesions and adjustment for age, ethnic group, smoking status and number of sexual partners.

Host and viral genes methylation and association with high grade cervical lesions by age.

There was a statistically significant interaction between age and methylation only in CADM1 gene (p-value of the interaction = 0.010). In particular for this gene the association with hgCIN risk was signifi- cantly higher in women over 35 years old [≥35 OR = 27.20 (95% CI 6.28–117.87) versus <35 OR = 3.14 (95% CI 1.50–6.58); p < 0.001 versus p = 0.002]. Similarly, methylation in MAL showed stronger associations in older women [≥35 OR = 19.84 (95% CI 5.82–67.62) versus <35 OR = 8.31 (95% CI 2.41–28.64); p < 0.001 versus p = 0.001]. For DAPK1 the association was not found.

For the three viral regions the associations in women under 35 years old were similar to those obtained in women over 35 years old (see Table 3 for details) except for L2 region.

The associations persisted also after adjustment.

Table 2
Risk of hgCIN by gene methylation.

	Co N	CIN2/3 N	Cut off	OR (95%CI)	p-Value[a]	Adj OR[b] (95%CI)	p-Value[a]	CIN2 N	OR (95%CI)	p-Value[a]	Adj OR[b] (95%CI)	p-Value[a]	CIN3 N	OR (95%CI)	p-Value[a]	Adj OR[b] (95%CI)	p-Value[a]
Methylation in host genes of all subjects																	
CADM1	294	247	5%	6.36 (3.37-11.91)	b0.001	6.52 (3.42-12.4)	b0.001	136	4.17 (2.03-8.56)	b0.001	4.34 (2.07-9.10)	b0.001	111	9.54 (4.80-18.97)	b0.001	9.21 (4.60-18.46)	b0.001
MAL	293	247	5%	12.45 (5.24-29.57)	b0.001	12.96 (5.38-31.24)	b0.001	135	5.98 (2.26-15.78)	b0.001	6.35 (2.33-17.30)	b0.001	112	22.66 (9.21-55.76)	b0.001	22.00 (8.80-54.99)	b0.001
DAPK1	289	214	9%	4.61 (1.48-14.34)	0.008	4.73 (1.51-15.11)	0.008	133	3.37 (0.93-12.13)	0.064	3.56 (0.96-13.23)	0.058	81	6.74 (1.92-23.64)	0.003	6.87 (1.92-24.57)	0.003
HPV+																	
CADM1	94	247	5%	2.77 (1.31-5.85)	0.008	2.31 (1.10-4.97)	0.033	136	1.82 (0.80-4.16)	0.154	1.65 (0.71-3.85)	0.245	111	4.17 (1.88-9.25)	b0.001	3.48 (1.53-7.90)	0.003
MAL	93	247	5%	4.58 (1.77-11.87)	0.002	3.63 (1.37-9.61)	0.009	135	2.20 (0.77-6.28)	0.141	1.94 (0.66-5.69)	0.225	112	8.34 (3.11-22.31)	b0.001	6.41 (2.33-17.59)	b0.001
DAPK1	91	214	9%	1.90 (0.53-6.82)	0.327	1.60 (0.43-5.93)	0.482	133	1.38 (0.34-5.69)	0.651	1.34 (0.32-5.66)	0.692	81	2.77 (0.69-11.11)	0.149	2.17 (0.50-9.42)	0.302
Methylation in viral regions for HPV types 16, 18, 31, 45																	
L1_I	28	117	5%	12.44 (3.81-40.63)	b0.001	12.49 (3.70-42.17)	b0.001	61	10.74 (2.66-43.31)	0.001	10.67 (2.60-43.85)	0.001	56	15.00 (3.00-74.98)	0.001	16.90 (3.13-91.31)	0.001
L1_II	26	110	5%	68.12 (8.16-568.40)	b0.001	58.44 (6.88-496)	b0.001	57	73.18 (4.07-1315.94)	0.004	59.28 (3.33-1054.54)	0.005	53	32.50 (3.86-273.65)	0.001	26.07 (3.00-226.27)	0.003
L2	26	114	5%	6.50 (2.34-18.06)	b0.001	6.37 (2.23-18.22)	0.001	60	4.73 (1.55-14.44)	0.006	4.92 (1.56-15.58)	0.007	54	10.62 (2.60-43.39)	0.001	9.35 (2.20-39.73)	0.002
HPV16																	
L1_I	15	98	5%	37.20 (9.17-150.96)	b0.001	–		48	30.00 (6.14-146.66)	b0.001	–		50	48.00 (8.13-283.42)	b0.001	–	
L1_II	12	91	5%	126 (12.88-1232.62)	b0.001	–		44	121.36 (6.06-2429.48)	0.002	–		47	64.4 (6.52-635.66)	b0.001	–	
L2	13	95	5%	21.5 (5.50-84.05)	b0.001	–		47	15.37 (3.58-65.96)	b0.001	–		48	33.75 (6.42-177.36)	b0.001	–	

Abbreviations: Co, controls; adj OR, adjusted odds ratio.
[a] p-Value assessed comparing controls versus cases CIN2/3.
[b] Adjustment for age, ethnic group, smoking status and number of partners.

3.2. Cumulative effect of host and viral genes methylation and high grade cervical lesions

The cumulative effect of the increasing number of host and viral methylated genes was considered as predictor of hgCIN (Table 4). A positive trend (p-trend<0.001) was found at increasing number of methylated human genes: one gene OR = 3.54 (95% CI 1.79–7.00); two genes OR = 6.48 (95% CI 2.48–17.00); three genes OR = 29.57 (95% CI 1.69–515.84).

As well, a positive trend (p-trend = 0.001) was found at increasing number of methylated viral genes. With all the three viral studied re- gions hypermethylated the risk of hgCIN strongly increased (OR = 36.54 (95% CI 3.95–337.00, p = 0.002)).

The associations also persisted in all the studied genes after adjustment.

4. Discussion

Our results showed that DNA methylation in HPV positive women is strongly associated with hgCIN confirmed by colposcopy. The study has been conducted in a setting involving a large Brazilian women popula- tion, with high number of HPV infection and hgCIN, recruited in the frame of awareness campaigns for adhesion to cervical screening, as in Brazil organized screening programs are not present.

At present, established cut off values for the assessment of the meth- ylation in host and viral genes are not available yet. In this study they were calculated by considering the limit of quantification of pyrose- quencing according to published indications [27-29] and for all the studied genes a cut off value not below 5% was obtained. Taking into ac- count that 5% is the widely recognized limit of detection of the pyrose- quencing to accurately discriminate methylation levels, the use of cut off values not inferior to 5% strengthens the reliability of our results. This conservative approach may at most lead to an underestimation of the association. Methylation was associated with high grade lesions either if it oc- curred in host or viral genes, with mean methylation values increasing with severity of the disease. The risk to develop hgCIN increased if the analyses were restricted to CIN3, being stable even after adjustment for confounders and confirming the hypothesis that CIN2 may be potentially a mixed diagnostic class, which may still gather a number of misclassified lesions. When the analyses were restricted to the HPV positive women, the association become statistically significant for all the studied regions, confirming the necessary role of HPV infection in challenging the initial steps of cervical cell transformation. Indeed, the mean level of methyla- tion in the host genes was lower in HPV negative women compared to HPV positive women without lesion. Altered methylation that potentially mediates progression probably occurs subsequently to HPV infection and likely with its persistence [3,4]. When considering CADM1 and MAL genes, our results are in line with previous reports which highlighted the role of their methylation in discriminating hgCIN [4,6,30,31], despite the use of different analytical methodologies (pyrosequencing or quanti- tative methylation specific-PCR). Association for DAPK1 resulted weaker, as also previously described by other authors [31].

Conversely, other authors who investigated methylation in the same regions that we targeted on the promoter of CADM1 and MAL genes did not find significant difference between women with normal and patho- logical cytology [16]. Since at present there is not a consensus about a gold standard method to investigate methylation status, it was argued that different methylation results can be obtained using different ana- lytical methods or targeting different CpG sites [3]. Otherwise, we could also speculate that different results could be obtained with differ- ent characteristics of the study population investigated, mostly due to screening history, sexual and lifestyle habit, and age. Although lifestyle habit may affect methylation status, we did not find effect on associa- tions, that did not varied after adjustment. When we stratified for age below and above 35 years old we found an increased risk of hgCIN in older women, stable after adjustment for confounders and particularly strong for hypermethylation in CADM1. In this unscreened population the persistence of virus and lesions may have favoured hypermethylation in host genes more in older than in younger women. The impact of

Table 3
Risk of hgCIN by gene methylation by age.

Age			Co	CIN2/3	OR (95%CI)	p-Value[a]	Adj OR[b] (95%CI)	p-Value[a]
Methylation of host genes								
<35	ALL	CADM1	185	157	3.14 (1.50-6.58)	0.002	3.08 (1.45-6.53)	0.003
		MAL	184	157	8.31 (2.41-28.64)	0.001	8.12 (2.32-28.39)	0.001
		DAPK1	181	134	4.93 (1.01-24.14)	0.049	5.00 (1.01-24.78)	0.049
	HPV+	CADM1	72	157	1.59 (0.68-3.70)	0.285	1.60 (0.68-3.75)	0.278
		MAL	71	157	3.12 (0.89-10.91)	0.075	3.20 (0.91-11.31)	0.070
		DAPK1	70	134	1.87 (0.38-9.27)	0.441	1.96 (0.39-9.75)	0.410
≥35	ALL	CADM1	109	89	27.20 (6.28-117.87)	<0.001	29.79 (6.72-132.03)	<0.001
		MAL	109	89	19.84 (5.82-67.62)	<0.001	22.18 (6.35-77.54)	<0.001
		DAPK1	108	79	4.36 (0.85-22.18)	0.076	4.99 (0.95-26.31)	0.058
	HPV+	CADM1	22	89	10.68 (1.37-83.25)	0.024	10.90 (1.37-86.83)	0.024
		MAL	22	89	5.61 (1.23-25.58)	0.026	5.88 (1.26-27.47)	0.024
		DAPK1	21	79	1.64 (0.19-14.45)	0.654	1.81 (0.20-16.48)	0.598
Methylation of HPV regions								
<35	HPV+	L1_I	21	70	11.17 (2.57-48.57)	0.001	12.01 (2.67-53.97)	0.001
		L1_II	20	67	44.00 (5.03-384.57)	0.001	47.81 (5.27-433.54)	0.001
		L2	20	68	8.40 (2.34-30.11)	0.001	8.33 (2.31-30.00)	0.001
	HPV+ 16	L1_I	12	57	25.20 (4.92-129.08)	<0.001	–	
		L1_II	10	54	79.50 (7.59-832.19)	<0.001	–	
		L2	11	55	22.31 (4.53-109.99)	<0.001	–	
≥35	HPV+	L1_I	7	47	16.87 (2.15-132.51)	0.007	14.48 (1.71-122.18)	0.014
		L1_II	6	43	48.33 (1.99-1171.25)	0.017	32.42 (0.98-1069.01)	0.051
		L2	6	46	4.10 (0.59-28.38)	0.153	2.67 (0.35-20.56)	0.346
	HPV+ 16	L1_I	3	41	110.6 (4.39-2787.89)	0.004	–	
		L1_II	2	37	75.00 (2.06-2724.83)	0.019	–	
		L2	2	40	32.27 (1.36-765.58)	0.032	–	

Abbreviations: Co, controls; adj OR, adjusted odds ratio.
[a] p-value assessed comparing controls versus cases CIN2/3.
[b] Adjustment for age, ethnic group, smoking status and number of partners.

Table 4
Risk of hgCIN by number of methylated genes.

	Number of host genes/viral regions with methylation above the cut-off	Co	CIN2/3	OR (95%CI)	p-Value[a]	Adj OR[b] (95%CI)	p-Value[a]	p-Trend[c]
Host genes								
ALL	0	270	155	Reference		Reference		
	1	13	27	3.54 (1.79-7.00)	b0.001	3.79 (1.90-7.56)	b0.001	
	2	5	20	6.48 (2.48-17.00)	b0.001	6.56 (2.46-7.56)	b0.001	
	3	0	8	29.57 (1.69-515.84)	0.002	29.10 (1.65-514.52)	0.021	b0.001
HPV+	0	77	155	Reference		Reference		
	1	9	27	1.44 (0.66-3.17)	0.361	1.22 (0.55-2.74)	0.625	
	2	4	20	2.27 (0.79-6.53)	0.128	1.79 (0.61-5.24)	0.289	
	3	0	8	8.47 (0.48-148.71)	0.144	5.91 (0.33-105.80)	0.228	0.073
HPV regions								
HPV+	0	5	1	reference		reference		
	1 or 2	7	10	7.14 (0.68-75.50)	0.102	6.85 (0.62-75.40)	0.116	
	3	13	95	36.54 (3.95-337.00)	0.002	32.78 (3.41-315.00)	0.003	0.001

Abbreviation: Co, controls; adj OR, adjusted odds ratio.
[a] p-Value assessed comparing controls versus cases CIN2/3.
[b] Adjustment for age, ethnic group, smoking status and number of partners.
[c] p-Trend = p-value obtained by considering the categorical variable as a continuous variable in the logistic model.

methylation on the risk of hgCIN was even more evident when viral re- gions were considered, and emphasized when HPV16 is involved. Methylation in each of the three selected regions showed a strong association with hgCIN, suggesting that viral methylation could be a predictive marker of hgCIN.

The cumulative effect of the host and viral genes methylation signif- icantly improved the association with risk of CIN2/3 lesions. Indeed detection of two or three methylated viral sites strongly increased the risk, as well the contemporary detection of more methylated host genes. At least for viral methylation, even the detection of only one methylated CpG site seemed sufficient to discriminate women at risk of disease pro- gression, also in the assumption that methylation events globally in- volve the HPV CpGs as an adaptive response of the cell to the external stimulus. This further evidence, although validation in wide cohorts with available follow up information is advisable, would suggest a pos- sible application of viral methylation analysis in clinical settings in the aim to select the women at higher risk of hgCIN.

The study had some limitations. First, follow up information are not available, thus limiting our investigation to a cross sectional evaluation. Second, the adherence to the campaign for adhesion to the cervical screening was voluntary. Recruitment of volunteers is a potential source of selection bias arising when volunteers from a specified sample may exhibit exposures or outcomes different from those of non-volunteers. However in this study all the comparisons and analyses performed were internal to the study population, therefore they were unlikely affected by selection bias. Third, we only considered four types of HPV (HPV16, 18, 31 and 45) for viral methylation analyses. Although these types were the most investigated in the recent literature and were en- countered among the most frequent in hgCIN and cervical cancer [13,32], extension of the analyses to other high risk types would allow to explore if implication of viral methylation could be independent from HPV type involved. We focused on single infections that, if persis- tent, were supposed to show higher methylation levels than multiple- type infections [13] and to be found in 90% of hgCIN [33].

To our knowledge this is the first study on the association among host and viral methylation and hgCIN involving a wide number of high grade cervical lesions. Association was found despite the methyla- tion differences observed between cases and controls were small, mainly due to abnormal cells mixed to exfoliated normal cervical cells in case samples.

In conclusion our data supported the role of methylation of both the host and the viral selected genes as a marker of aggressiveness in cervi- cal lesions. Furthermore our results showed that is possible to assess the methylation status using the scrape sample obtained at enrolment contributing to a future translational application of these molecular analyses.

AcknowledgThents

We acknowledge the "Vale per la vita" Onlus Association for the precious support to our research through laboratory instrument donation.

Financial acknowledgThent

This work was supported by Italian Ministry of Health (GR-2011- 02349870); Alliance FUNPAR-LIGH.

Conflict of interest

The authors declare that they have no competing interests.

References

[1] T. Wittenberger, S. Sleigh, D. Reisel, M. Zikan, B. Wahl, M. Alunni-Fabbroni, et al., DNA methylation markers for early detection of women's cancer: promise and chal- lenges, Epigenomics 6 (2014) 311–327.
[2] P.A. Jones, J.P. Issa, S. Baylin, Targeting the cancer epigenome for therapy, Nat. Rev. Genet. 17 (2016) 630-641.
[3] M.A. Clarke, P. Luhn, J.C. Gage, C. Bodelon, S.T. Dunn, J. Walker, et al., Discovery and validation of candidate host DNA methylation markers for detection of cervical precancer and cancer, Int. J. Cancer 141 (2017) 701-710.
[4] A.T. Lorincz, Virtues and weaknesses of DNA methylation as a test for cervical cancer prevention, Acta Cytol. 60 (2016) 501–502.
[5] M. Clarke, N. Wentzensen, L. Mirabello, A. Ghosh, S. Wacholder, A. Harari, et al., Human papillomavirus DNA methylation as a potential biomarker for cervical can- cer, Cancer Epidemiol. Biomark. Prev. 21 (2012) 2125–2137.
[6] N. Wentzensen, M.E. Sherman, M. Schiffman, S.S. Wang, Utility of methylation markers in cervical cancer early detection: appraisal of the state-of-the-science, Gynecol. Oncol. 112 (2009) 293-299.
[7] N. Wentzensen, M. Schiffman, T. Palmer, M. Arbyn, Triage of HPV positive women in cervical cancer screening, J. Clin. Virol. 76 (2016) S49–S55.

[8] M.L. Tornesello, L. Buonaguro, P. Giorgi-Rossi, F.M. Buonaguro, Viral and cellular bio- markers in the diagnosis of cervical intraepithelial neoplasia and cancer, Biomed. Res. Int. 2013 (2013) 519619.
[9] K. Louvanto, E.L. Franco, A.V. Ramanakumar, N. Vasiljević, D. Scibior-Bentkowska, A. Koushik, et al., Methylation of viral and host genes and severity of cervical lesions associated with human papillomavirus type 16, Int. J. Cancer 136 (2015) (E638- E634).
[10] P. Liu, M. Iden, S. Fye, Y.W. Huang, E. Hopp, C. Chu, et al., Targeted, deep sequencing reveals full methylation profiles of multiple HPV types and potential biomarkers for cervical cancer progression, Cancer Epidemiol. Biomark. Prev. 26 (2017) 642-650.
[11] K. Hoelzer, L.A. Shackelton, C.R. Parrish, Presence and role of cytosine methylation in DNA viruses of animals, Nucleic Acids Res. 36 (2008) 2825–2837.
[12] N. Vasiljević, D. Scibior-Bentkowska, A. Brentnall, J. Cuzick, A. Lorincz, A comparison of methylation levels in HPV18, HPV31 and HPV33 genomes reveals similar associ- ations with cervical precancers, J. Clin. Virol. 59 (2014) 61–66.
[13] N. Wentzensen, C. Sun, A. Ghosh, W. Kinney, L. Mirabello, S. Wacholder, et al., Meth- ylation of HPV18, HPV31, and HPV45 genomes is associated with cervical intraepithelial neoplasia grade 3, J. Natl. Cancer Inst. 104 (2012) 1738–1749.
[14] L. Mirabello, M. Schiffman, A. Ghosh, A.C. Rodriguez, N. Vasiljevic, N. Wentzensen, et al., Elevated methylation of HPV16 DNA is associated with the development of high grade cervical intraepithelial neoplasia, Int. J. Cancer 132 (2013) 1412–1422.
[15] L. Mirabello, C. Sun, A. Ghosh, A.C. Rodriguez, M. Schiffman, N. Wentzensen, et al., Methylation of human papillomavirus type 16 genome and risk of cervical precancer in a Costa Rican population, J. Natl. Cancer Inst. 104 (2012) 1–10.
[16] C. Sun, L.L. Reimers, R.D. Burk, Methylation of HPV16 genome CpG sites is associated with cervix precancer and cancer, Gynecol. Oncol. 121 (2011) 59–63.
[17] J.L. Brandsma, Y. Sun, P.M. Lizardi, D.P. Tuck, D. Zelterman, G.K. Haines 3rd, et al., Dis- tinct human papillomavirus type 16 methylomes in cervical cells at different stages of premalignancy, Virology 389 (2009) 100-107.
[18] L. Marongiu, A. Godi, J.V. Parry, S. Beddows, Human Papillomavirus 16, 18, 31 and 45 viral load, integration and methylation status stratified by cervical disease stage, BMC Cancer 14 (2014) 384.
[19] M. Kalantari, K. Osann, I.E. Calleja-Macias, S. Kim, B. Yan, S. Jordan, et al., Methylation of human papillomavirus 16, 18, 31, and 45 L2 and L1 genes and the cellular DAPK gene: considerations for use as biomarkers of the progression of cervical neoplasia, Virology 448 (2014) 314-321.
[20] I. Murakami, T. Fujii, K. Dan, M. Saito, A. Ohno, T. Iwata, et al., Methylation of human papillomavirus-52 and -58 is a candidate biomarker in cervical neoplasia, J. Clin. Virol. 58 (2013) 149–154.
[21] M.A. Clarke, A. Gradissimo, M. Schiffman, J. Lam, C.C. Sollecito, B. Fetterman, et al., Human papillomavirus DNA methylation as a biomarker for cervical precancer: con- sistency across 12 genotypes and potential impact on management of HPV-positive women, Clin. Cancer Res. 24 (2018) 2194–2202.
[22] A. Gillio Tos, M. Dag, V. Bicalho, C. Fiano, V. Grasso, L. De Marco Tarallo, et al., Case- control study of HLA-G promoter methylation status, HPV infection and cervical neoplasia in Curitiba, Brazil: a pilot analysis, BMC Cancer 12 (2012) 618.
[23] B. Izzi, A.M. Binder, K.B. Michels, Pyrosequencing evaluation of widely available bi- sulfite conversion methods: considerations for application, Med. Epigenet. 2 (2014) 28–36.
[24] C. Grasso, M. Trevisan, V. Fiano, V. Tarallo, L. De Marco, C. Sacerdote, et al., Perfor- mance of different analytical software packages in quantification of DNA methyla- tion by pyrosequencing, PLoS One 11 (2016), e0150483.
[25] J. Dunn, A. Baborie, F. Alam, K. Joyce, M. Moxham, R. Sibson, et al., Extent of MGMT promoter methylation correlates with outcome in glioblastomas given temozolo- mide and radiotherapy, Br. J. Cancer 101 (2009) (124–13).
[26] A.B. Håvik, P. Brandal, H. Honne, H.S. Dahlback, D. Scheie, M. Hektoen, et al., MGMT promoter methylation in gliomas-assessment by pyrosequencing and quantitative methylation-specific PCR, J. Transl. Med. 10 (2012) 36.
[27] U. Lehmann, Quantitative validation and quality control of pyrosequencing® assays, Methods Mol. Biol. 1315 (2015) 39–46.
[28] F. Busato, E. Dejeux, H. El Abdalaoui, I.G. Gut, J. Tost, Quantitative DNA Methylation analysis at single-nucleotide resolution by pyrosequencing®, Methods Mol. Biol. 1708 (2018) 427-445.
[29] T. Mikeska, J. Felsberg, C.A. Hewitt, A. Dobrovic, Analysing DNA methylation using bisulphite pyrosequencing, Methods Mol. Biol. 79 (2011) 33–53.
[30] A.T. Hesselink, D.A. Heideman, R.D. Steenbergen, M. Gök, F.J. van Kemenade, S.M. Wilting, et al., Methylation marker analysis of self-sampled cervico-vaginal lavage specimens to triage high-risk HPV-positive women for colposcopy, Int. J. Cancer 35 (2014) 880–886.
[31] N. Vasiljević, D. Scibior-Bentkowska, A.R. Brentnall, J. Cuzick, A.T. Lorincz, Credential- ing of DNA methylation assays for human genes as diagnostic biomarkers of cervical intraepithelial neoplasia in high-risk HPV positive women, Gynecol. Oncol. 132 (2014) 709-714.
[32] S. de Sanjose, W.G. Quint, L. Alemany, D.T. Geraets, J.E. Klaustermeier, B. Lloveras, S. Tous, et al., Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study, Lancet Oncol. 11 (2010) 1048–1056.
[33] J. van der Marel, J. Berkhof, J. Ordi, A. Torné, M. Del Pino, R. van Baars, et al., Attrib- uting oncogenic human papillomavirus genotypes to high-grade cervical neoplasia: which type causes the lesion? Am. J. Surg. Pathol. 39 (2015) 496-504.
[34] S.R. Giolo, J.M.P. Soler, S.C. Greenway, M.A.A. Almeida, M. de Andrade, J.G. Seidman, et al., Brazilian urban population genetic structure reveals a high degree of admix- ture, Eur. J. Hum. Genet. 20 (2012) 111–116.